NASA-TM-112380

New technology

Surgical Endoscopy

Surg Endosc (1994) 8: 817–820

This section is designed to bring forward some of the latest innovative technology with explanations in terms that will clarify their importance to the discipline of surgery. Through the efforts of the Innovative Technology Committee of the Society of American Gastrointestinal Endoscopic Surgeons (SAGES), leading experts in various areas will be invited to present a summary of new technology, often including their pioneering work.

NDB

7N-54-TM

036 267

Human interface technology

An essential tool for the modern surgeon

R. M. Satava,[1,2] S. R. Ellis[3]

[1] Walter Reed Army Medical Center, Washington, D.C. 20307, USA
[2] Advanced Research Projects Agency (ARPA), 3701 North Fairfax Drive, Arlington, VA 22203, USA
[3] NASA Ames Research Center, Moffett Field, CA 94035, USA

Received: 10 May 1993/Accepted: 17 December 1993

Summary. Human interface technology is a new science which must be understood by all surgeons in order to cope with the ever-increasing complexity of surgical practice. This science is the understanding of how humans comprehend, interact, and use the world around them. The increasing use of robotics, computers, and virtual reality depend upon this technology to create a "user-friendly" environment to be able to assimilate the massive amount of data and images and to "naturally" interact with machines and computers. Through careful implementation, more complex systems will become easier to use and enhance the surgeon—the technology must adapt to the surgeon, not the reverse.

Key words: Robotics — Telepresence — Virtual reality — Human interface technology — Ergonomics

Until recently, surgeons operated directly upon patients, touching and cutting the tissues. With the advent of laparoscopic surgery and other endoscopic procedures, the surgeon no longer directly sees or touches the organs or tissues which he removes. Instead he watches an electronically transmitted video image (television monitor) and inserts long, awkward instruments into hidden body cavities to extract or repair organs. This is one giant step forward for the patient and two giant steps backward for the surgeon. The reason is that this new technology evolved as an extension and modification of previously unrelated technologies, tied to traditions and designs of the past, without regard to natural or intuitive interaction between the surgeon and the machines, computers, or instruments which he is using. The classic example is laparoscopic surgery, where surgeons are willing to accept loss of stereoscopic vision, near-complete loss of sensory "feel" of the tissue, and extremely awkward instruments for dissection and excision. Laparoscopic surgery just "happened" — surgeons began using whatever equipment was available, and as new instruments were designed, they were simply modifications of the same basic instruments rather than a redesign to take into consideration the use and interaction of the surgeon with the video image, computers, or instruments. Since video imaging and computers will continue to play an increasingly prominent role in surgical procedures and promise to considerably reduce their cost, we must begin to include an analysis of human performance capabilities, with all the biological limitations and idiosyncrasies, into the design of the new surgical systems so as to augment rather than impede the surgeon's abilities. The science of understanding and integrating man into these advanced technologies is called human interface technology (HIT).

Definitions

Human interface technology enhances a person's abilities while performing a task which requires a tool or synthetic device; it "empower[s] the user" [4] by providing an understanding of the interaction of man with a tool, machine, computer, or other instrument in the real world. In order to understand the field of HIT, definitions are needed to advance the basic concepts. The *environment* is the place where the interaction

The opinions or assertions contained herein are the private views of the author and are not to be construed as official, or as reflecting the views of the Department of the Army of the Department of Defense.

Correspondence to: R. M. Satava, ARPA-DIRO, 3701 N. Fairfax Drive, Arlington, VA 22203, USA

between a person and the world occurs—"The theater of human activity" [2]. For HIT applications there are two types of environments—*real* and *synthetic*. The *real environment* includes the real world and all the actions which we perform on a day-to-day basis—in particular, its content, its geometry, and its rules of interaction or dynamics. A *synthetic environment* generally refers to a computerized graphic drawing with mathematically contrived rules and dynamics in which a person interacts. (See below.) Either the real or the synthetic environment can be the subject matter for the application of human interface technology.

The expression *virtual reality* (VR) has come into common usage recently and the term is the source for some confusion and scientific debate. Most commonly, VR refers to a *synthetic environment;* it is an imaginary "world" that may be created as a simulation within a computer and which changes viewpoint with the natural movement of the user's head. There are many alternative names, such as *artificial reality* and *electronic reality;* however, none of these descriptions adds scientific value above the designation of synthetic environments. Since synthetic environments are "within" the computer, they are all *virtual environments*, and if the realism and accuracy are high enough the users may be convinced that they have entered an imagery world, the one that only exists inside the simulation computer. It is like *Alice and Wonderland*, where you can climb down the rabbit hole and enter an illusory, interactive, apparently real world as if it actually existed. Typical examples of such an illusion are found in modern flight simulators, though some elements are also present in video games using 3-D format.

History

Some of the earliest attempts to understand HIT grew out of the fighter pilot and space programs, where pilots and astronauts required sophisticated equipment to perform the highly complex tasks in the aviation environment. Attempts to solve these problems by incorporating robots and remote manipulators revealed complete absence of scientific data on how man interacted with machines. Studies began in "human factors engineering," "man–machine interface," ergonomics, and other fields in order to better design tools to enhance human productivity. The first successful results were produced in flight simulators, and then in the middle 1960s, Ivan Sutherland [6] developed the first helmet-generated object that he manipulated in three dimensions. With the addition of a tracking device on the head and hand, and of electronic glove (such as a DataGlove from VPL Research in 1985) instead of a joystick, this system evolved into the more familiar current-day VR systems. By this means, the person is now "empowered" to act upon the objects in the imaginary, virtual world as if they actually existed, using the DataGlove to pick up a teapot, turn a light switch on and off, cut with scissors, and so forth. The interface, the HMD and glove, are potentially natural and intuitive; you are totally immersed in an imaginary world; when you look around you see things as they are in a real world; and when you grab an object it can be picked up. They way you manipulate objects is so natural that you don't think you are giving commands to a computer when you are pointing your finger. Thus the interface could be considered "transparent" because you do not see or think about the commands you give to the computer (you don't type "pick up"), you simply do it.

Research in stereoscopic vision has become intense since the introduction of the HMD. There are many perceptual and cognition issues which are totally unknown, though most research to date has indicated that 3-D vision greatly enhances performance during manipulative tasks. Although it is not possible to go into depth on these critical aspects of 3-D vision, it suffices to emphasize that 3-D vision must be considered whenever HIT analysis occurs.

Concepts and application

To put together a successful, fully integrated interface which is "transparent" to the person performing a task, many factors must be taken into consideration. Since the majority of the newer solutions for tasks use and are dependent upon computers, discussion will focus around enhancing computer and manipulator technology to augment human capabilities. It is important to differentiate between robotics and telepresence. Robotics have much in common with artificial intelligence (AI): the "robot" is designed to have an intelligence of its own and to replace the human; but telepresence is based on intelligence amplification (IA), enhancing a human's own abilities [1]. Hence the controversy in control theory of AI vs IA; in the latter the control is by the human; you keep "the man in the loop." For example, to pick up an object using robotics, the person pushes the "pick up" control button and the robot performs the task; using telepresence, the person uses a joystick, electronic glove, or other "input device" and moves the remote arm precisely with his hand to pick up the object. The human interface to a purely robotic system is principally restricted to discrete inputs (such as pushing buttons or typing commands), although supervisory control of a robot might use more advanced computer-graphic interfaces. In contrast, displays for teleoperation, telepresence, and virtual environments are intimately involved with all aspects of HIT and critically depend upon good design of the interface for success.

The key to HIT is that it must be intuitive; it should mimic as closely as possible the natural way a human would perform a task. We use all of our senses as we go about our daily business; therefore, the more sensory input given back to the person in a coordinated fashion, the more realistic the interface. In current-day telepresence there is visual, auditory, tactile, and force feedback. More importantly, not only is there multisensory reception by the person, but the senses are presented in their natural spatial configuration.

That is, the stereophonic sound comes exactly from the position of the object in the environment producing the sound, and the eye–hand axis between the real hand or tool and the eyes is maintained. One reason laparoscopic surgery is difficult is because of loss of the eye–hand axis: During open surgery the surgeon looks in a single axis from his eyes directly through his hand at the scissors and then cuts; laparoscopically there are two axes, and the surgeon looks up at the monitor while trying to cut below the line of vision. The oculovestibular apparatus in the head is sending "up" signals to the brain, while the proprioception and kinesthetic senses in the hands are sending "down." This is a conflict of sensory input, and thus the intuitive reflexes must be unlearned, and a new set of subliminal actions must be learned. This is one of the main reasons surgeons require training in laparoscopic surgery and cannot just begin doing it.

The way visual information is displayed is crucial to the intuitiveness of a system and not necessarily the same for all applications. An HMD is used to give the feeling that you are "within" the imaginary world and that you can move around the entire environment freely; this is referred to as an "immersive" environment. Looking into a monitor or at a wall screen is called a "through the window" or nonimmersive environment. Any approach to virtual reality or telepresence must be either one or the other type—they are mutually exclusive. For example, with open surgery, the surgeon stands or sits next to the patient and "looks down" onto the operating field; therefore in telepresence surgery (see below) a through-the-window approach is chosen so the surgeon "looks down into" the video monitor just as if he were at the operating table. Since a surgeon does not have the perspective of being *inside* of the patient during open surgery, an HMD is not an intuitive interface for telepresence surgery. If an HMD display were chosen, there should be a compelling advantage to offset the loss of intuitiveness. Another example of selecting the display to meet the task exists at the University of North Carolina, where architects use virtual reality to "walk through" their virtual buildings; a treadmill in front of a wall-size monitor is used (rather than an HMD) to recreate the illusion of actually walking through the building. The choosing of an appropriate display is referred to as *metaphor*—using the right symbol or context (e.g., HMD, monitor, wall screen) to mimic how a person would actually view and interact with that particular environment. The most widely known computer metaphor is the *desktop* of Macintosh computers, which represents the computer screen as if it were the top of an actual desk. As a matter of completeness, it is important to note that there are many other forms of visual displays. The HMD could be "see through" (transparent) with a computer-generated display (virtual world) superimposed upon the real world; or it could be a miniaturized 2-inch display placed just off the center of the field of view such as the Private Eye or Virtual Vision, to provide information. Other options include large-screen displays, or video walls or domes, as in flight simulation.

Not only must the output from the tool or computer to the person (e.g., tactile sensation, video display, etc.) be intuitive; the way the surgeon gives input back through the tool or to a computer also is important. In virtual reality, a DataGlove is a very good example of a potentially natural "input device" since many of the gestures that you make with your hand (and "read" by the glove) are intuitive. When you want to move in a virtual world, you point your finger (as if to say "Go there"), or to pick up something you make a fist as if grabbing the object. Unfortunately, although the pointing gesture is intuitive, it is not very accurate, and it results in a movement with the sensation of "flying" rather than walking through the environment. Also, prolonged use of the glove is tiring, and it can interfere with using the hand for manipulation. This demonstrates limitations and tradeoffs with the use of the glove as an input device. This technique of task analysis must be rigorously applied whenever considering usage of any input/output device or display for long, complicated surgical cases.

For an input interface in telepresence surgery, the actual handles of surgical instruments are used, so the surgeon grabs the same handles that he would use in open surgery and moves the instruments in precisely the same fashion as in open surgery. In laparoscopic surgery the instruments are designed to work around a fulcrum. Thus, when using the instrument, most of the manipulation is done from the shoulder in a "rowing motion" rather than from the hand or wrist, which is much more precise, delicate, and similar to open surgery. The HIT rule that the input device should feel and be used as naturally as possible is the reason telepresence surgery input was designed to mimic open surgery and truly does represent an improvement upon laparoscopic surgery.

Other critical components in real and virtual environments are tracking devices, such as head trackers, eye trackers, etc. These instruments provide the feedback to the computer as to where the person is located in relation to the real or virtual world. A Polhemus position tracker is a magnetic device which can be located on any part of the body; typically one on the head or HMD will give the proper orientation of the head and one on the hand allows for gesture input. It is these sensors that "map" the person to the world and permit coordinated interaction.

An example and interesting HIT application of trackers is to have the laparoscopic camera remotely controlled, thereby freeing up a camera assistant. Using a telemanipulating hand to hold the camera, control by the surgeon is possible via a foot pedal, a button, a joystick, voice command, or a tracking device. To determine which would be best, observation of a surgeon during open surgery (task analysis) reveals that to obtain a better or different view, the surgeon simply moves his head left, right, in, out, up, or down; in laparoscopic surgery he might tell the camera operator assistant to move in a certain direction. Thus the a priori solution using HIT methodology for control would be a head (or eye) tracker or a voice-activated system. An additional benefit of either of these two methods is that they are "unencumbered" input devices; that is, they leave the surgeon's hands free to continue the procedure without interruption. Of interest is the fact that a tracking device is totally "transparent"; in all of daily life situations, when a person wants a better view of an object or procedure, they reflexively move their head (or refocus their eyes) without exerting a conscious effort to detract their attention.

The concepts of visual displays, input devices, trackers, and sensors are all key components of HIT. The following are two examples of current application of HIT to advanced-technology systems: The Green telepresence surgery system [3] and the virtual-reality surgical simulator [5]. Both are early prototypes and are in the demonstration phase.

The telepresence surgery system consists of two components—the operative site and the surgical workstation. At the operative site, there is a manipulator, paired video cameras, and a stereophonic microphone; at the surgical workstation there is the 3-D monitor, hand controllers, and stereophonic speakers. Employing HIT methods, the surgical workstation (which is the interface between the surgeon and the system) was designed to imitate open surgery. Therefore, the surgeon looks "down" into the 3-D monitor and grasps handles of surgical instruments that are in the same eye–hand axis relationship as open surgery. In addition, the handles provide force-feedback sensory input and have precisely the same motion as operating with open surgery—there is no fulcrum effect. A surprising finding has been that the addition of the stereophonic sound significantly enhances the feeling of presence, emphasizing the previous statement that the realism of an event is dependent upon the number of different sensory inputs. Thus, the surgeon is given the illusion that what is being operated upon actually exists in front of him/her.

The virtual-reality surgical simulator provides a computer-generated abdomen with simulated organs and surgical instruments (of cartoon-level fidelity) and the typical HMD and DataGlove. The interaction is in real time such that the surgeon's motion corresponds to the actions; when an instrument is picked up it appears in the hands and when the clamp is squeezed, it grasps the tissue. Some of the HIT involved in the surgical simulator corresponds well to open surgery, such as grasping objects and clamping tissue; however, surgeons do not operate with helmets, and in that sense, the HMD may be a poor choice for an interface.

Discussion

As the surgical world becomes more technologically complex, it is essential that we design new surgical systems that are intuitive and have no learning curve imposed upon the surgeon—the surgeon must not learn the tool; the tool must accommodate the surgeon. The primary step is to incorporate HIT at the beginning of the development of any new surgical tool or system. This is done by identifying the environment in which the task occurs and by a critical analysis of the task or procedure in the natural environment with the purpose of completely understanding how the surgeon would perform the task without the system. Then the new tool or system must be designed and used in such a fashion that the system is transparent, that the surgeon is totally unaware of the system and believes he is doing the task in the way that is natural. The proof of the success of any system will be the amount of learning or retraining that will be required to begin using the new system. The ideal HIT will be one with no learning curve, one in which the surgeon will be able to begin using a new tool, procedure, or system without special training. As the technological complexity of the surgical arena increases, we must use HIT to leverage the technology to provide a simpler environment for the surgeon.

References

1. Brooks FC. Personal communication
2. Ellis SR (1991) Pictorial communication in virtual and real environments (prologue). Taylor and Francis, Philadelphia, pp 3–11
3. Green PE, Piantanida TA, Hill JW, Simon IB, Satava RM (1991) Telepresence: dexterous procedures in a virtual operating field (abstr). Am Surg 57: 192
4. Laurel B (1990) The art of human-computer interface design. Addison-Wesley Publications, Menlo Park, CA, pp xi–xii
5. Satava RM (1993) Virtual reality surgical simulator: the first steps. Surg Endosc 7: 203–205
6. Sutherland IE (1965) The ultimate display. Int Fed Info Process 2: 506